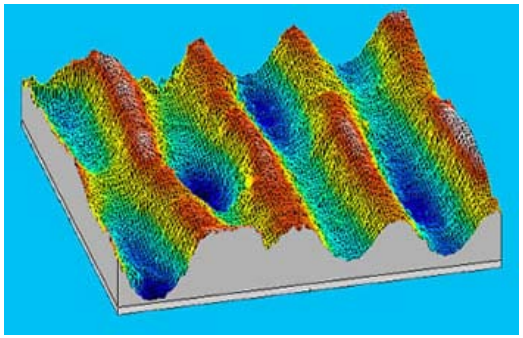
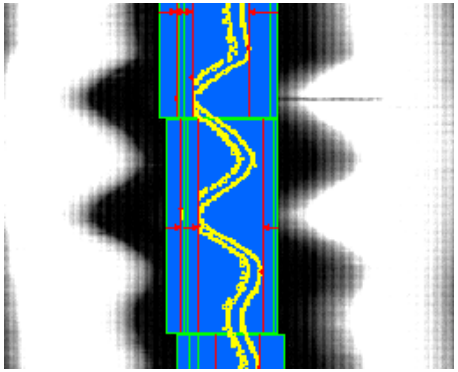




Optical Scanning Applied to Recorded Sound Preservation and Access

Carl Haber

Lawrence Berkeley National Lab

Collaboration and Support

Lawrence Berkeley National Lab: Earl Cornell, CH, Vitaliy Fadeyev, Robert Nordmeyer, Mitch Golden + students: Maryrose Barrios, Nicolas Scozzoro, Wei Zhou (2010)
Library of Congress: Peter Alyea, Larry Appelbaum, Dianne van der Reyden, Elmer Eusman, Eric Hansen, Gene DeAnna
EIF Fribourg – Switzerland: Ottar Johnsen + students: Adrian Nicolet, Tobias Mueller (2010)

INSTITUTE of Museum and Library SERVICES

NATIONAL ENDOWMENT FOR THE HUMANITIES

THE ANDREW W. MELLON FOUNDATION

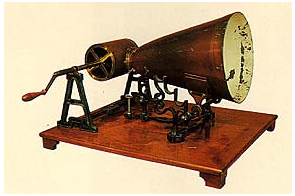
John Simon Guggenheim Memorial Foundation
Fellowships to Assist Research and Artistic Creation

America's Collections

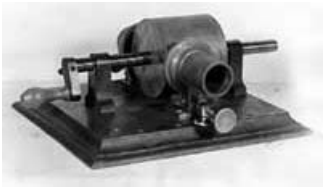
Fig. 4.1 U.S. Institutions Have Taken Responsibility to Preserve 4.8 Billion Collections Items	
Books/Bound Volumes	1.7 billion
Microfilm/Microfiche	1 billion
Natural Science Specimens	820.2 million
Photographic Collections	727.4 million
Archaeological Collections, individually cataloged	197.8 million
Unbound Sheets, cataloged in items	95.8 million
Online Files	54.6 million
Historic Objects	48.3 million
Recorded Sound	46.4 million
Unbound Sheets, cataloged in linear feet	43.6 million
Moving Images	40.2 million
Art Objects	21.2 million
Digital Materials	9 million
Archaeological Collections, cataloged in cubic feet	2.6 million

Heritage Health Index 2005

- US museums, libraries, and institutions collectively hold 4.8 Billion items
- 46.4 million sound recordings
  - 9.6 million are in grooved formats
  - 59% of these are in unknown condition
  - And there is the rest of the world
- Archivists want to reformat all pre-digital media since playback systems are not maintained
- **National Recording Preservation Act of 2000** "A bill to…maintain and preserve sound recordings and collections of sound recordings that are culturally, historically, or aesthetically significant…, " (Public Law 106-474; H.R.4846).



History



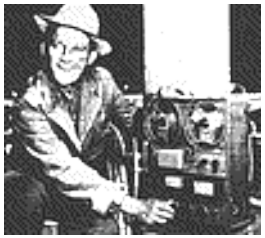
- 1853 Leon Scott: *Phonoautograph* paper recorder
- 1877 Thomas Edison invents sound reproduction on tin foil *Phonograph*

- 1880-5 Bell(s) and Tainter, Volta Lab research into audio formats, finally introduce wax cylinder
- 1887 Emile Berliner invents disc *Gramophone*
- 1925 Western Electric *Orthophonic* (electrical) system

end of the "**Acoustic Era**"

Discos fonográficos Pathé
Caras y Caretas (7/7/1906)

- 1929 Edison production ends, lacquer transcription disc introduced
- 1947 Magnetic tape in production use, Ampex 200A
- 1948 33 1/3 rpm LP introduced
- 1958 Stereophonic LP on sale, uses 45/45 system



- 1963 Cassette magnetic tapes

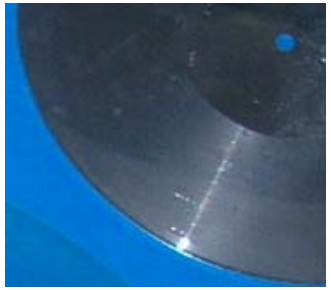
- 1982 Compact Disc (*CD*)

end of the "**Analog Era**"

- 2001 Apple *IPOD*
- Late 2000's Massive online access to digital sound files

Sound Carrier Diversity and Condition

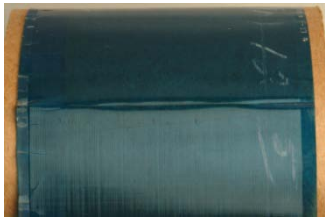
Wax cylinders
Fungus
Wax bloom



Al disc: oxidation

Acetate: flaking

exudation



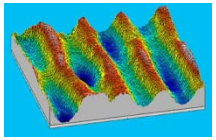
Plastic belts
breakage

dirt

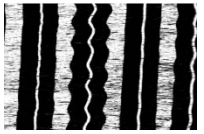
warpage

The Berkeley Optical Sound Restoration Project

- We are developing methods of optically restoring mechanical sound recordings **without contact** to the medium – like text scanning
- Address concerns of the preservation, archival, and research communities:
  - Preservation: Restore or stabilize delicate or damaged media
  - Access: Mass digitization of diverse media, automation
  - Assessment
  - Obsolete formats and legacy playback systems
- The approach evolved naturally out of methods of optical metrology, pattern recognition, image processing, and data analysis we use for physics research.

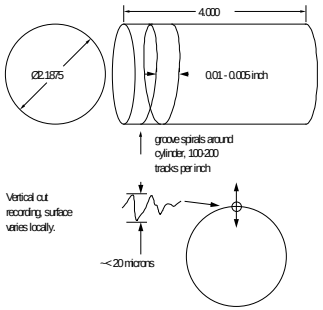


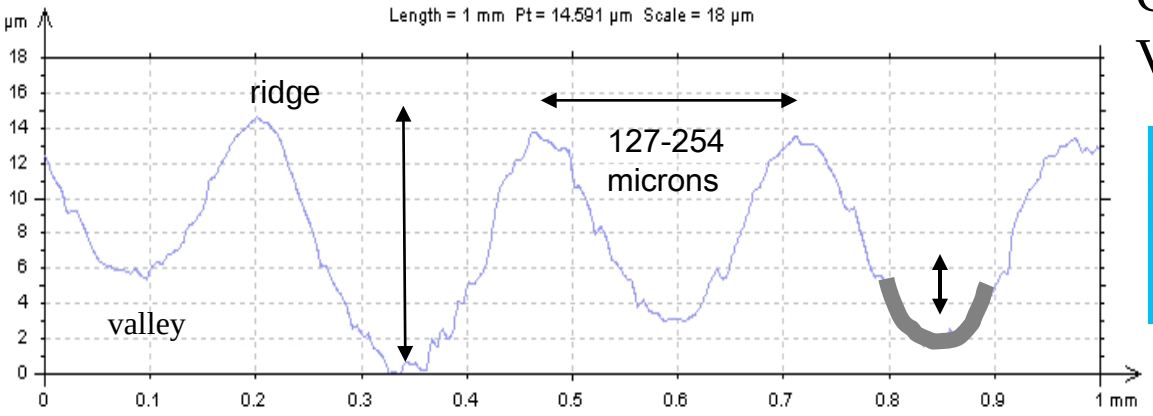
Non-Contact Digital Imaging



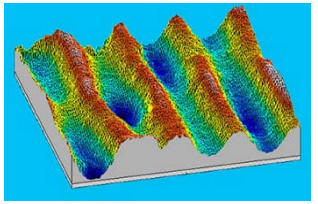
- Create a high resolution digital map of entire surface
- Computer plays record (image) with a virtual stylus
- Product
  - Standard digital sound files (ie .wav)
  - High resolution digital images which may be reanalyzed later as well
- Protects samples from further damage
- Repair existing damage through “touch-up”
- Reconstruct broken records
- Offload aspects of restoration to automated software

A “smart” copying machine for records



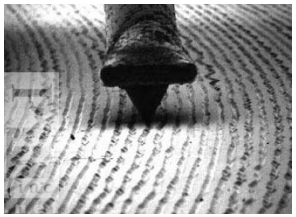


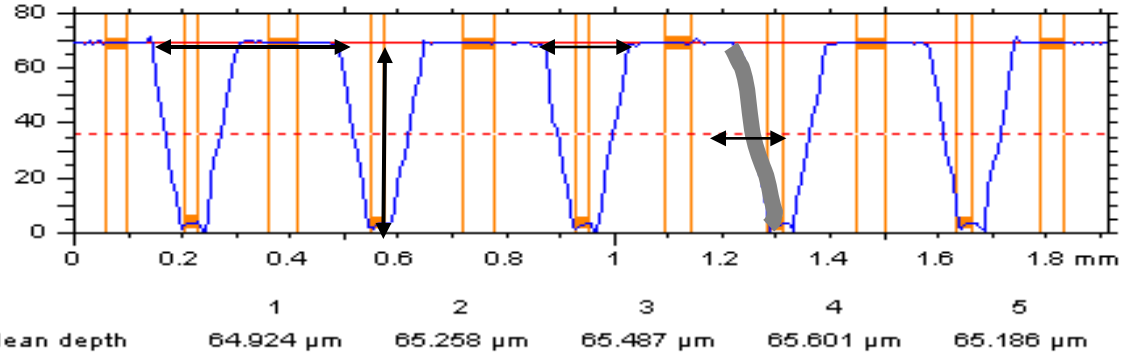
Cylinder surface
Vertical cut



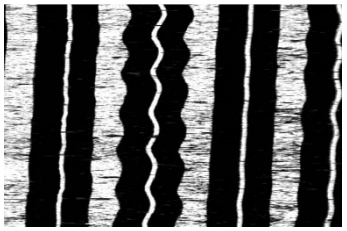
3D image

400 µm 150 µm





Disc surface
Lateral cut

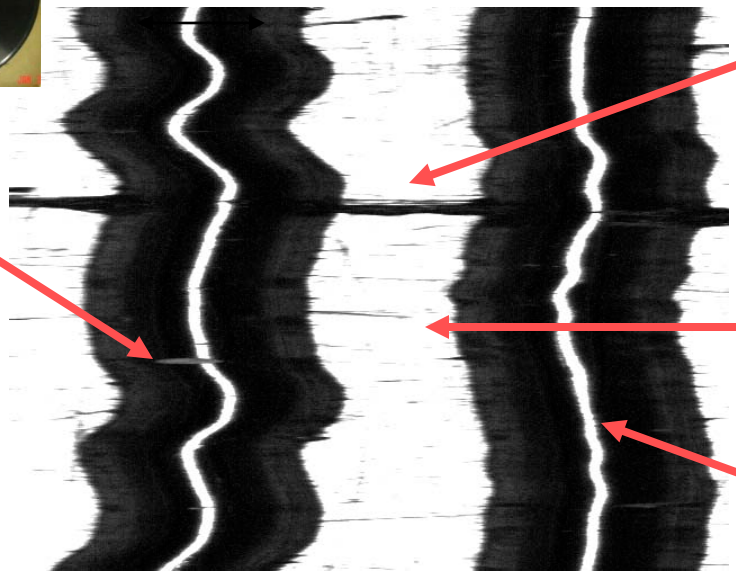


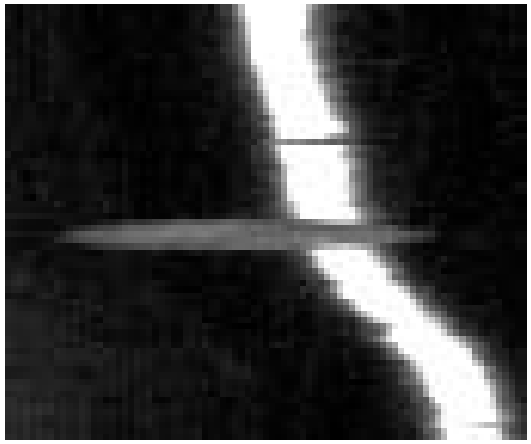
2D image

Audio is encoded in micron scale features which are >100 meters long.

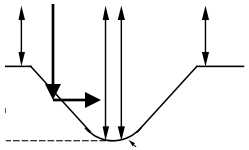
2D Imaging: Electronic Camera

160 μm

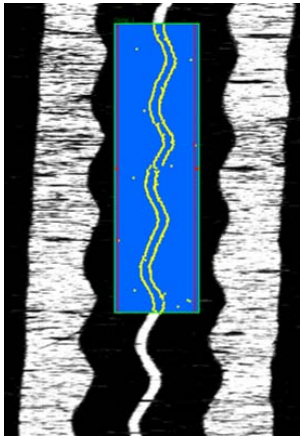




Coaxial illumination

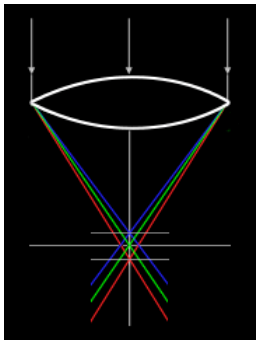


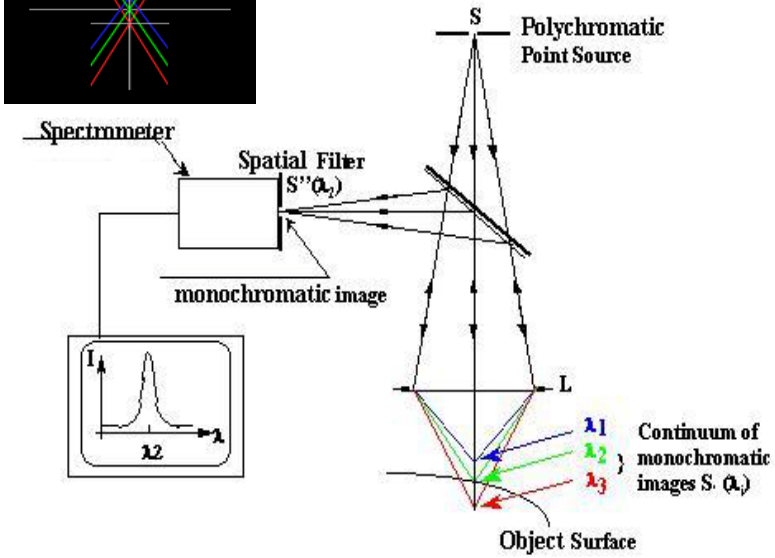
- Suitable for disc with lateral groove
- Require 1 pixel = ~ 1 micron on the disc surface
- High resolution yeilds narrow depth of field, 10 – 20 microns
- High speed cameras allow near “real-time” imaging
- Extract groove information from high contrast edge transitions

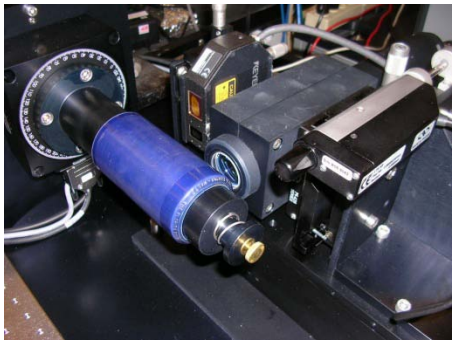


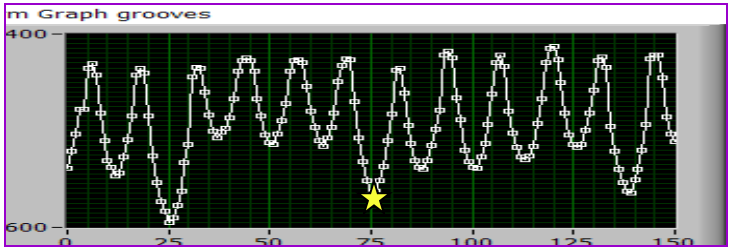
3D Imaging: Confocal Scanning Probe

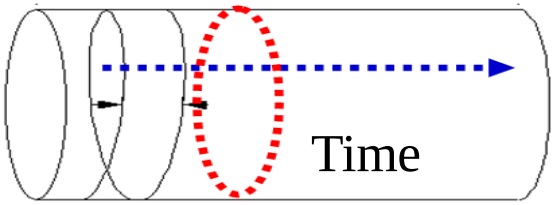
Required for cylinder with vertical groove modulation.

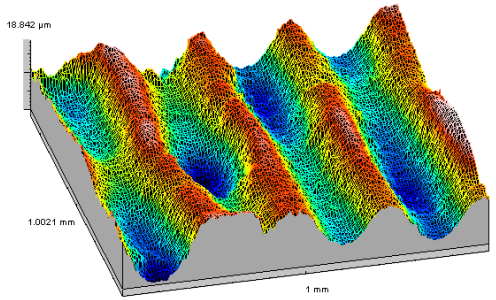












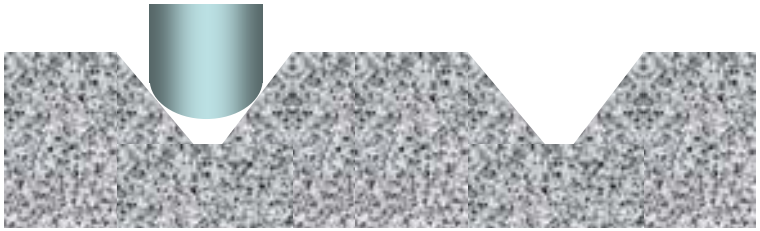
1 point at 2 KHz vs. 180 points at 1.8 KHz

3 days 30 minutes

Redundancy

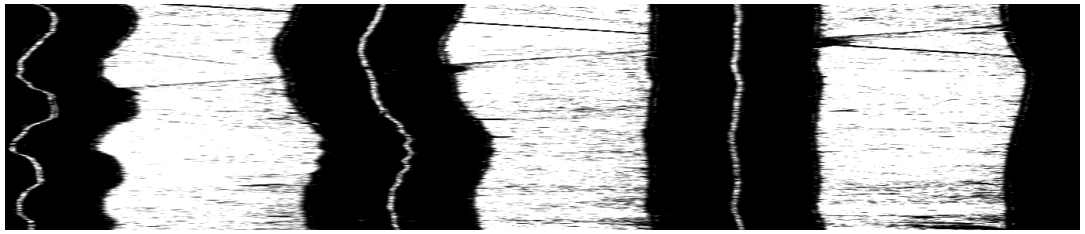
Audio stored in entire profile, signal averaging

Stylus



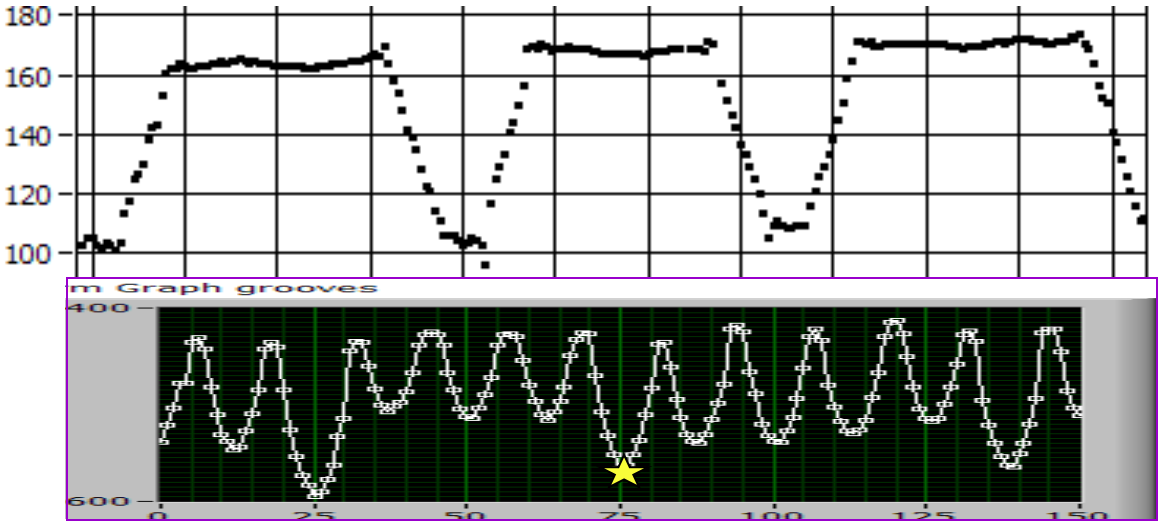
Stylus comes into contact at 2 points

2D



Measure the groove bottom with 2 points/slice

3D



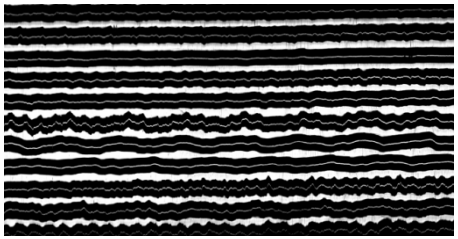
Measure the entire groove with ~30 points/slice

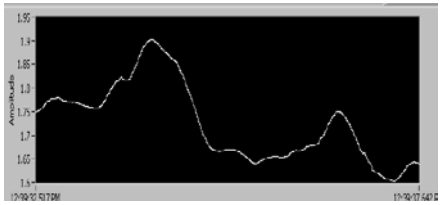
Developments

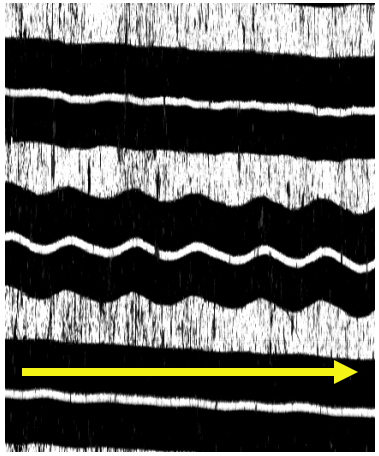
- Concept was tested 2002-2003 leading to interest and support from the Library of Congress and others.
- IRENE: a fast 2D optical scanner for disc records 2006-7 (NEH)
  - Digital access to the most common media + special formats
  - Installed 2006 at Library of Congress, evaluation, upgrade
- 3D: a fast 3D scanner for cylinders and discs 2008-9 (IMLS)
  - Preservation and restoration of early and damaged recordings
  - Benefit from recent improvements in 3D probe technology
- Connecting to Collections: 2010-12 (IMLS)
  - Migration of technology into use at multiple collection sites
  - Special Studies:
    - Wax field recorded and dictation cylinders
    - Damaged broken, unplayable, or tare recordings
    - Early experimental recordings
    - Cylinder molds and disc stampers

I.R.E.N.E

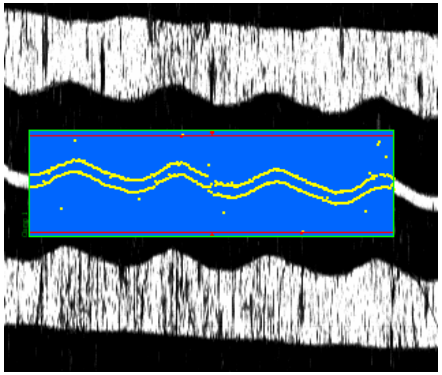
Line scan



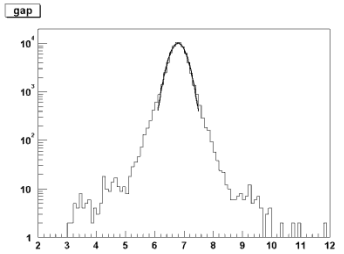




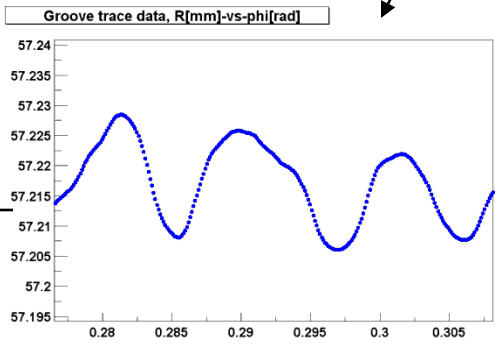
Time
Pixels = 104 KHz

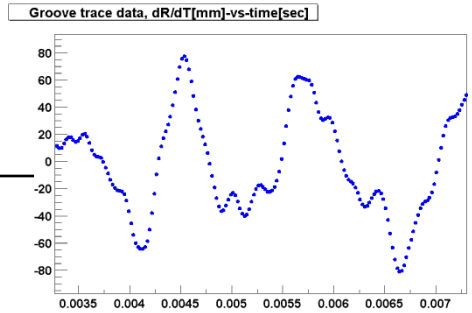


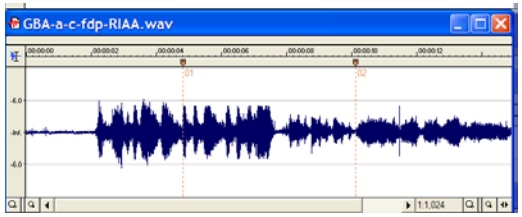
Width across
groove bottom



Measure slope
at each point
(stylus velocity)

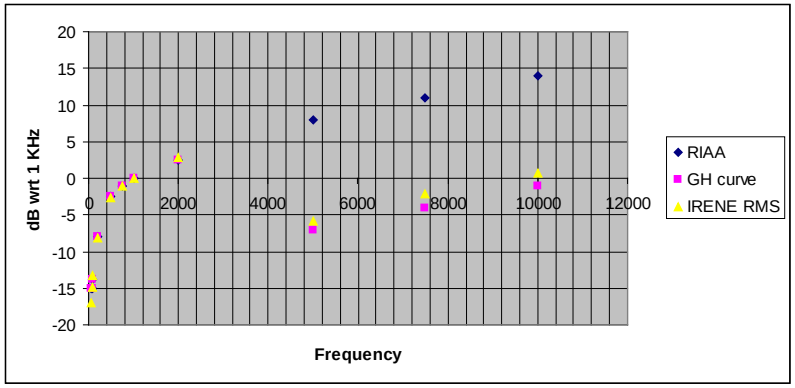






Average
Filter using width cut

Pristine test disc

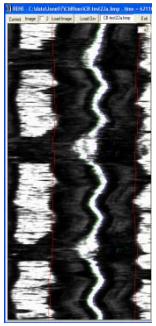


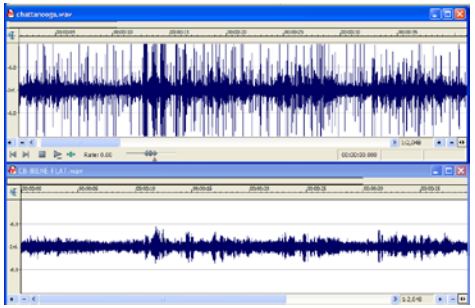
Les Paul 1953

Stylus-skip

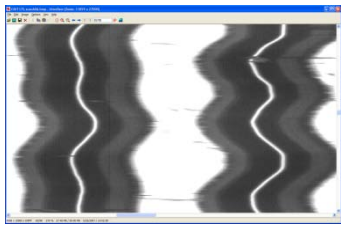
IRENE

Ida Cox 1923 heavy wear





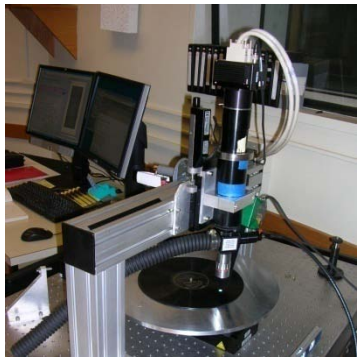
Broken disc
1930

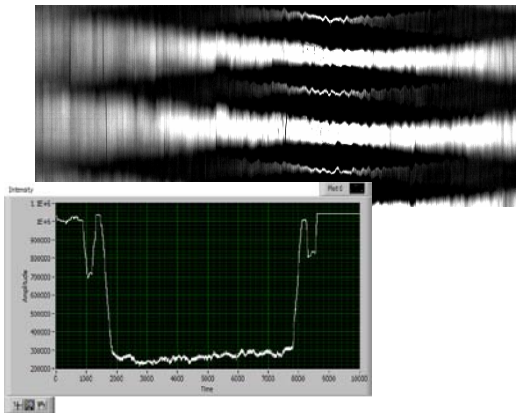


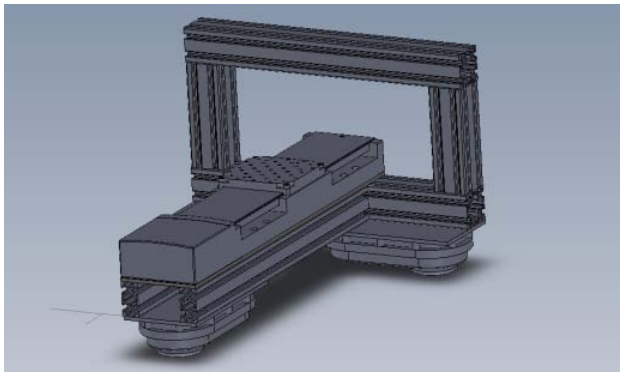
Lacquer radio transcription
1950

Collections Scanning

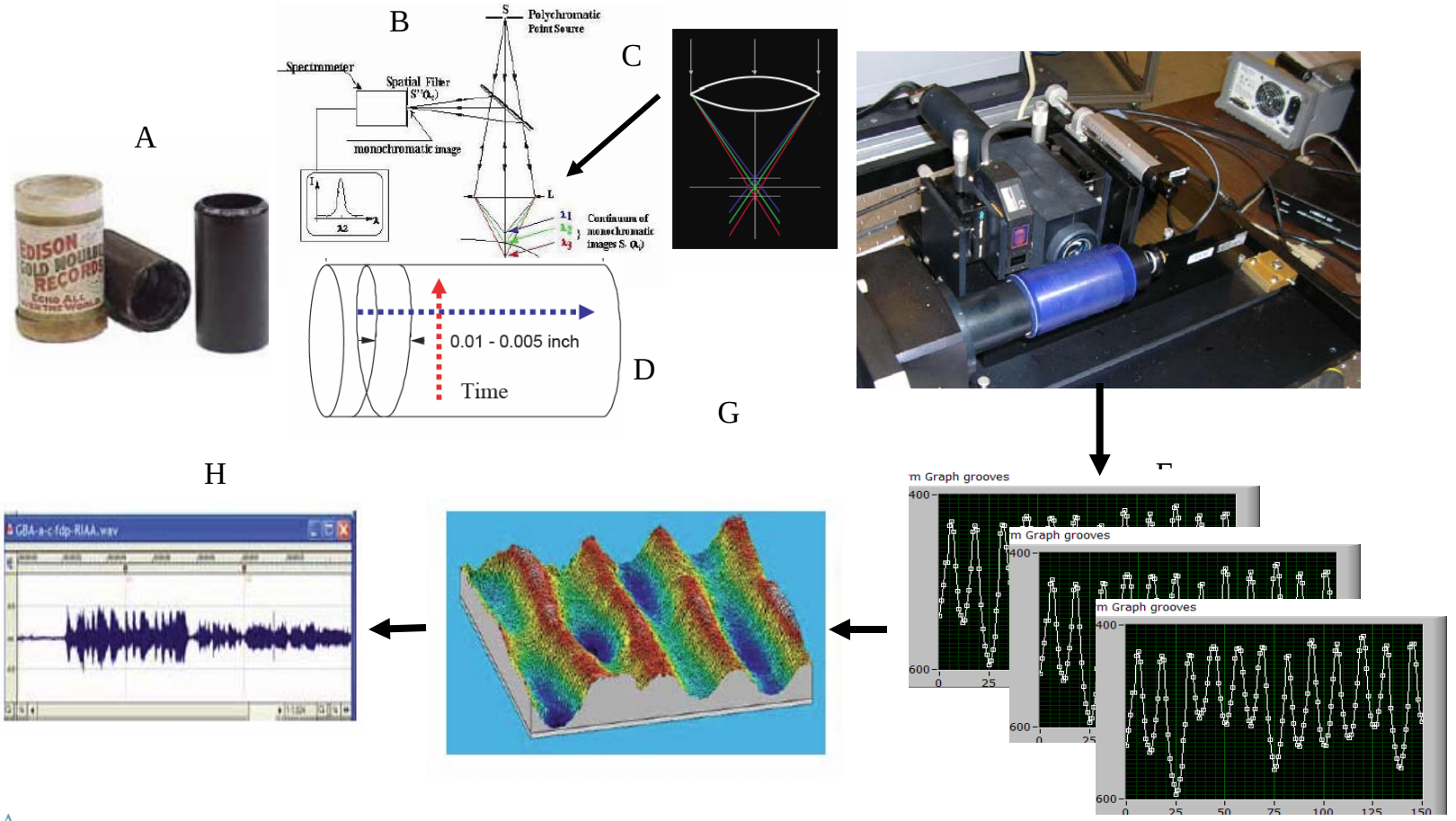
- An IRENE system is installed at NAVCC-Culpeper, Va.
  - Has been upgraded towards a production application but it's use is still considered a study in progress
  - Fast setup, database access, logging
- A "portable" version is being built for the University of Chicago South Asia Library for use in Chennai, India



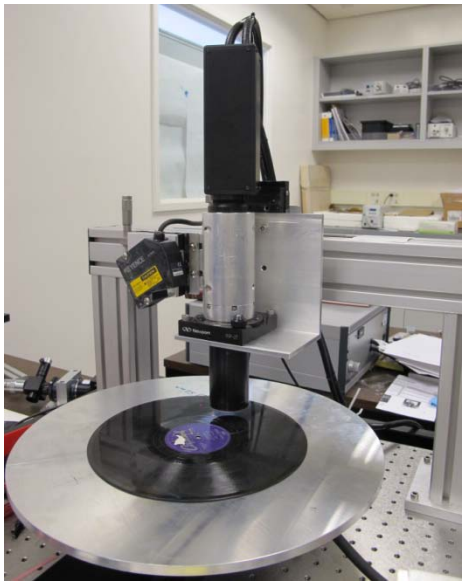


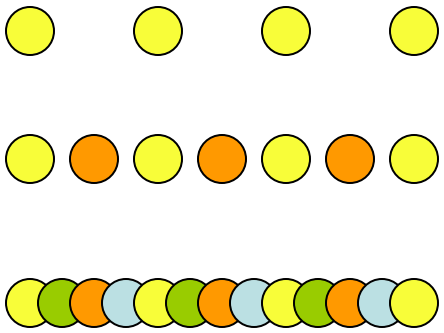


Basic 3D Scanning Process



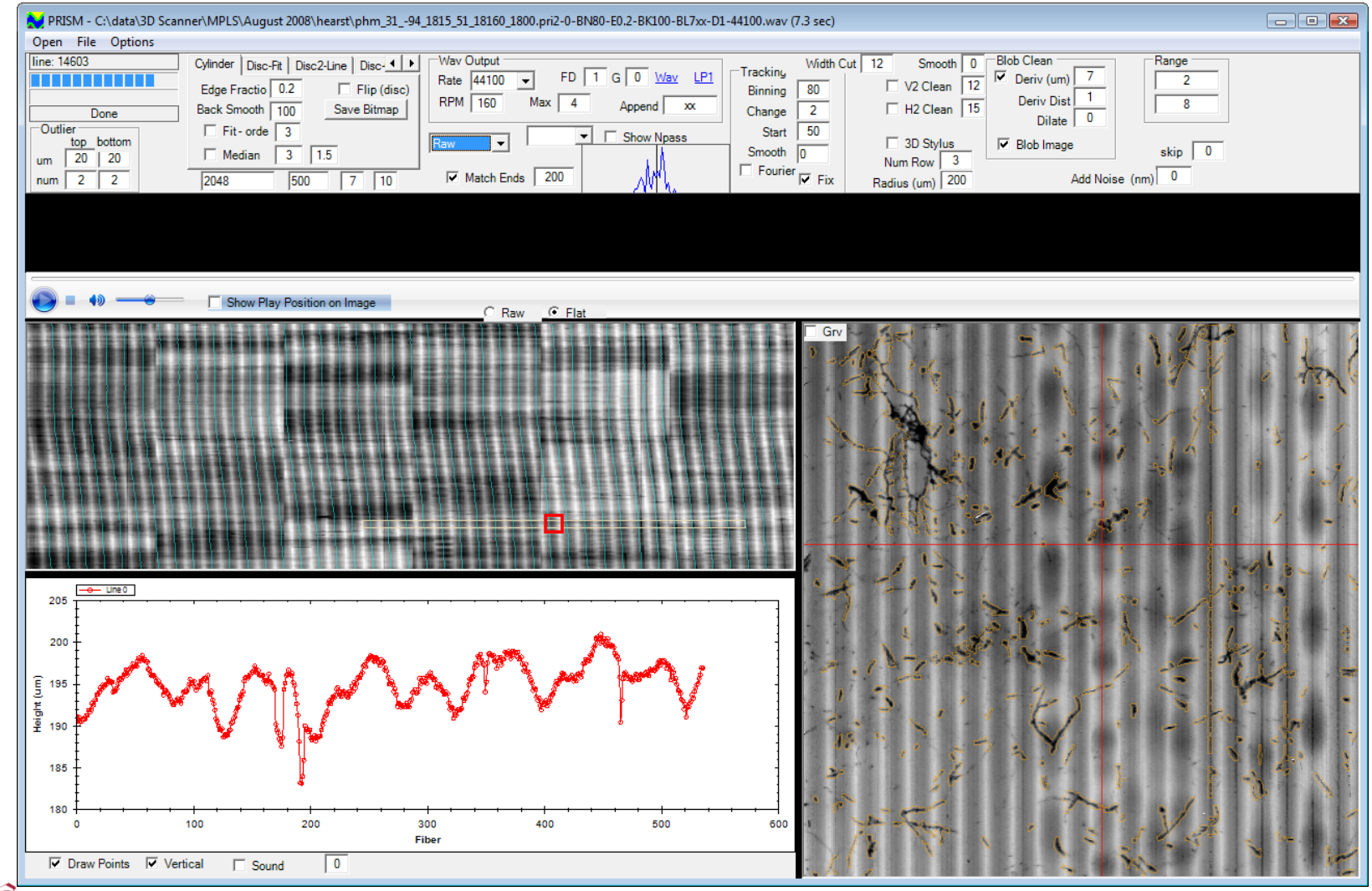
Fast 3D Scanner Project





- Funded 10/07 by IMLS
- Build fast scanner using new 180 point confocal probes
- Integrate into simple and robust control and analysis system.
- Installed at the Library of Congress in 2010

3D Analysis Software

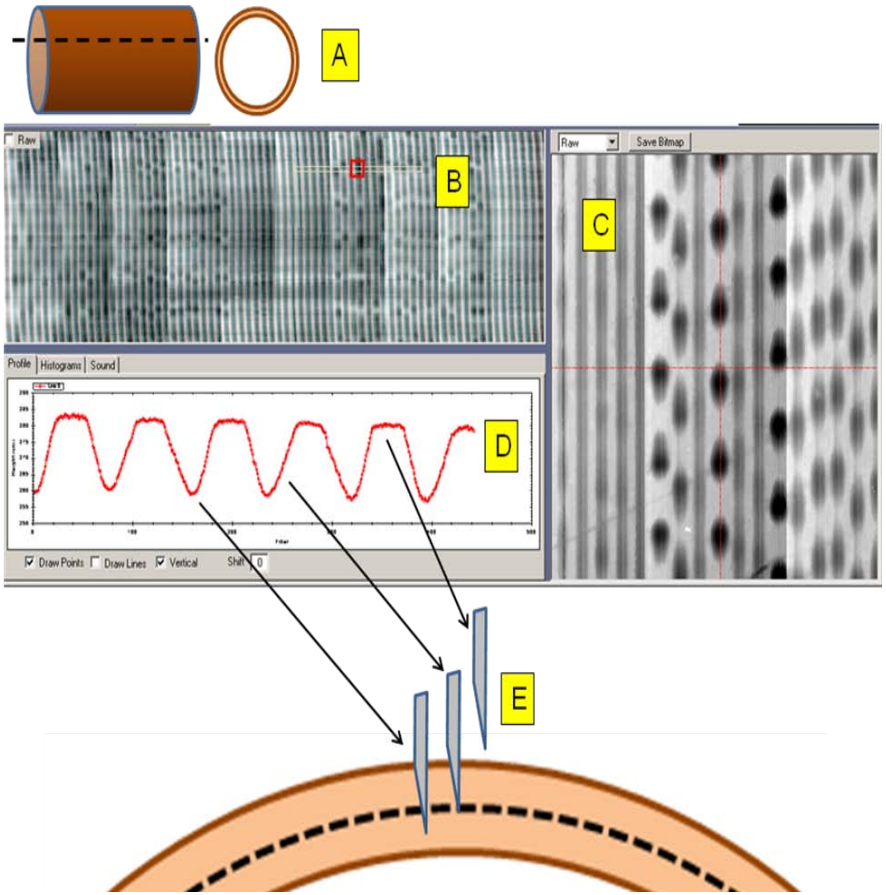


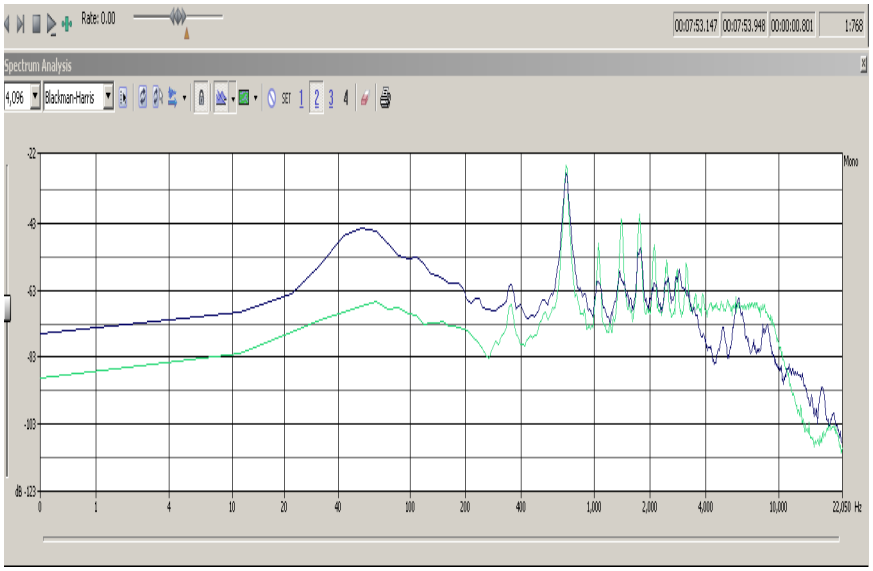
Special Studies

- Examples which challenge the technology
- Expanded study on wax field cylinders from ethnographic collections
  - Lots of surface deflects due to fibers, cutting process, filtering and cleanup
- Early experimental recordings
  - A range of unusual materials, formats, and methods
- Berliner discs
- Copper galvanos at the Berlin Phonogramm Archive
  - Key international collection, unique application of confocal scanning
- Broken cylinders and discs
  - Fixture and s/w development with EIF Fribourg, ENHS Dickson
- Rare transcription media
  - Lacquer, aluminum, plastic (recording sessions, fieldwork, novelty, Presidential)

Field Recordings Pilot Study

- 1930 Vancouver Island, Franz Boas, audio+film
- UW Burke Museum, IU-Archive of Trad. Music, AMNH
- Systematic scan of 12 cylinders in a variety of conditions
- Reconstruction using PRISM with derived parameters
- Compare one and four pass, stylus (1959, 1984)





Stylus version (black)

3D version (blue)

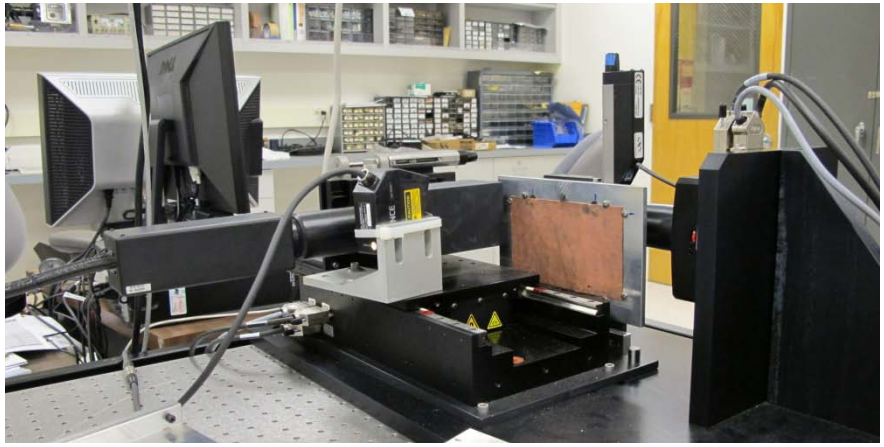
Since optical scanning is free from the real-time dynamic effects inherent in stylus playback certain types of distortion can be reduced

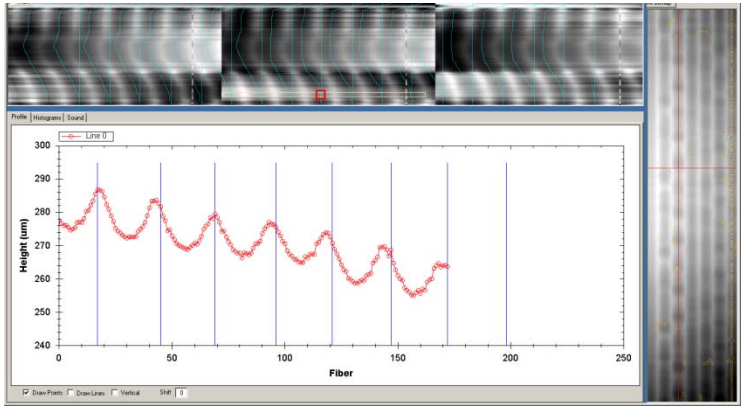
Galvano Study

- Berlin Phonogramm Archive, large scale conversion of field recordings to galvanos as a means for preservation and access
- Many casts have not been made, old technology
- Earlier attempts to scan with 2D video
- Proposal to make a thin probe to fit inside galvano
- Proof of principle using sacrificial sample

Stylus

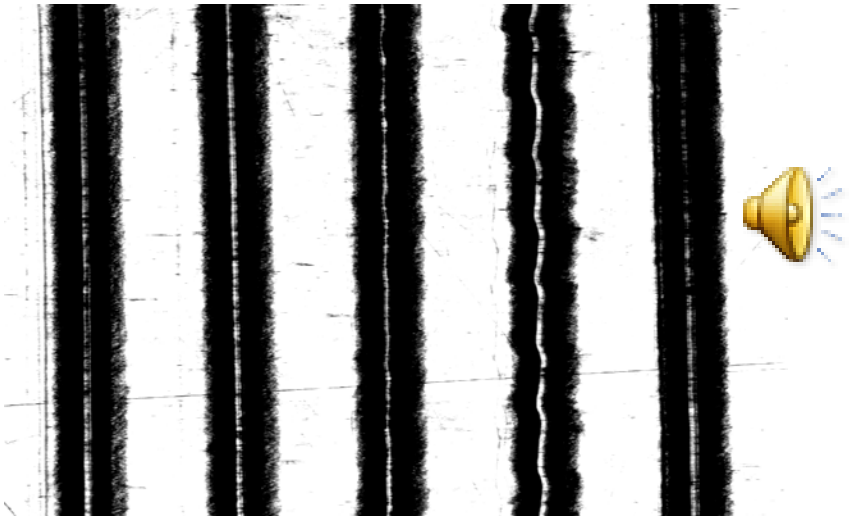
3D

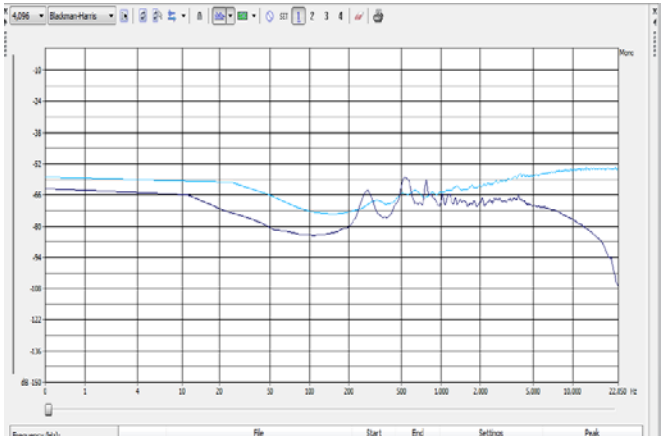


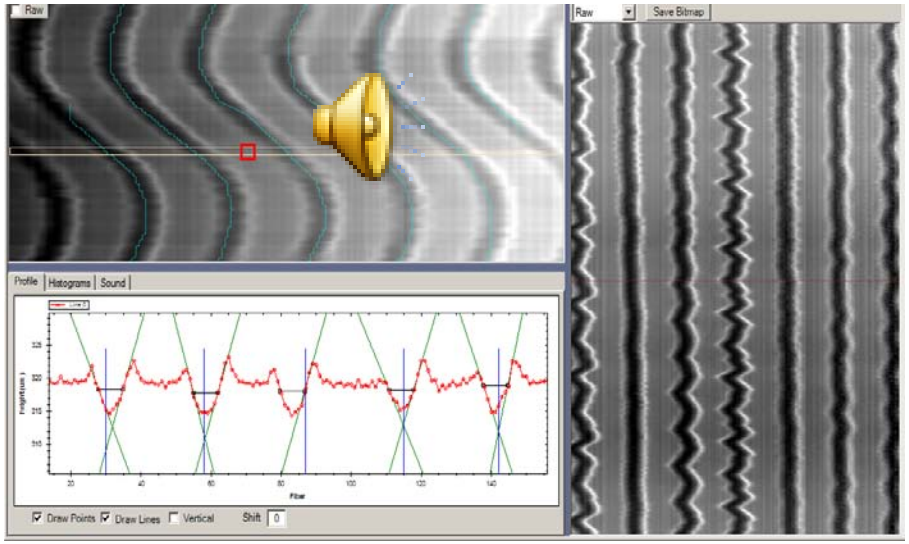


Aluminum Disc Records

- Study on discs from Harvard Millman Parry South Slavic fieldwork collection
- Comparison of IRENE and 3D imaging

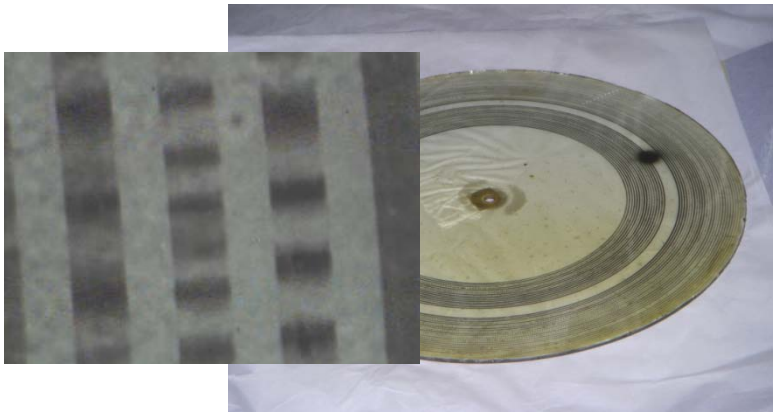






Experimental Recordings

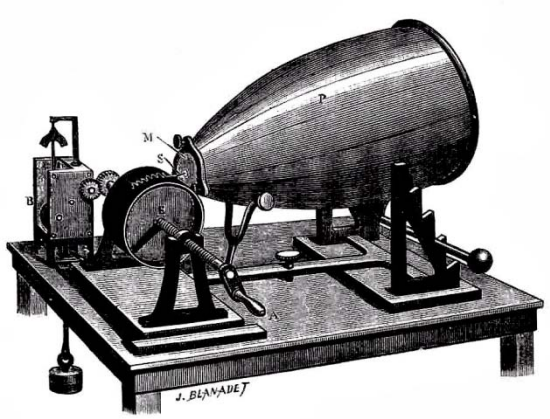
- Diverse media from the development of audio recording technology
- Leon Scott phonoautograms ~1860, restored with IRENE analysis methods
- A.G. Bell, C.S. Tainter, C.Bell, 1880-1888, Volta Lab – SI NMAH collection, a broad R&D study which led to the adoption of wax as a recording medium
- Both cylindrical and disc formats were studied before Berliner, using mechanical and photographic sound carriers
- Near term goal is to scan sample of the Bell collection as a pilot study

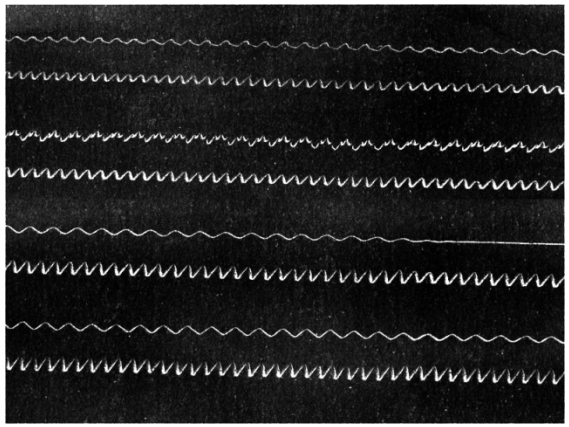


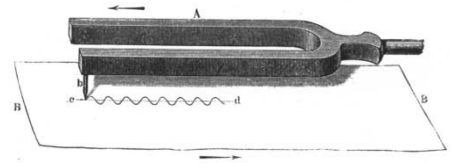
Early Recorded Sound

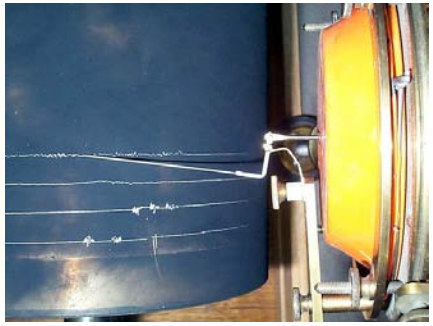
Phonautograph
Leon Scott
1853

Scott enscribed sound on paper and could not play it back



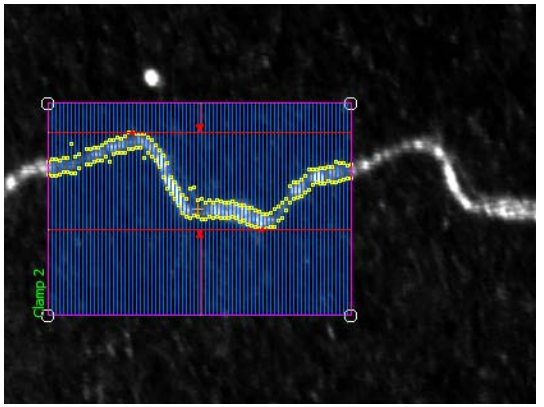


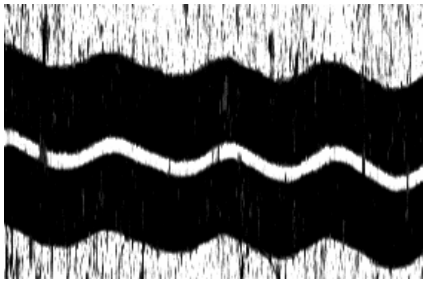


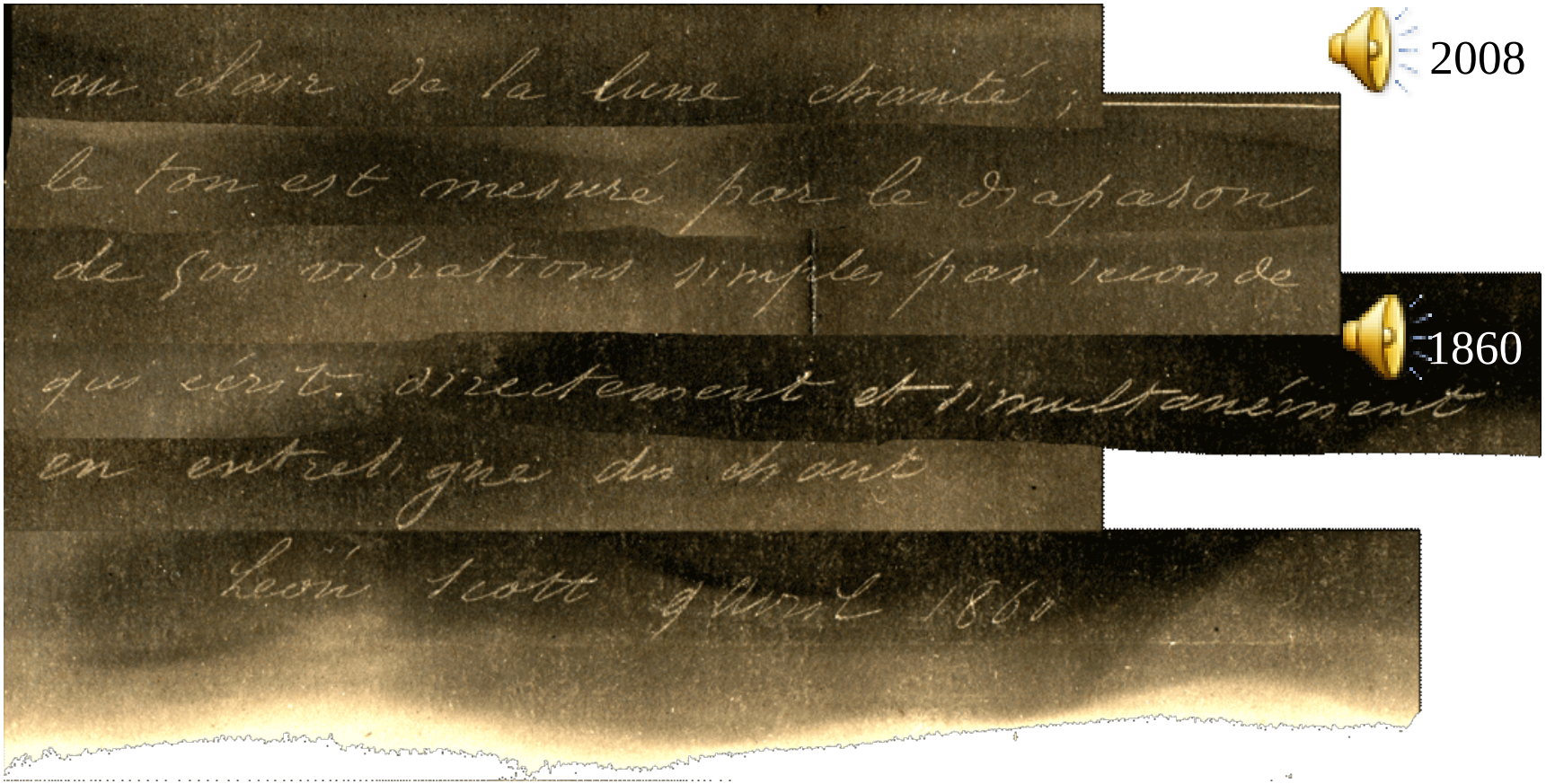


Recorded April 9, 1860
Deposited in the French Academy of Sciences

…visually similar to "IRENE" 2D scans, can processed and analyzed by the same tools…

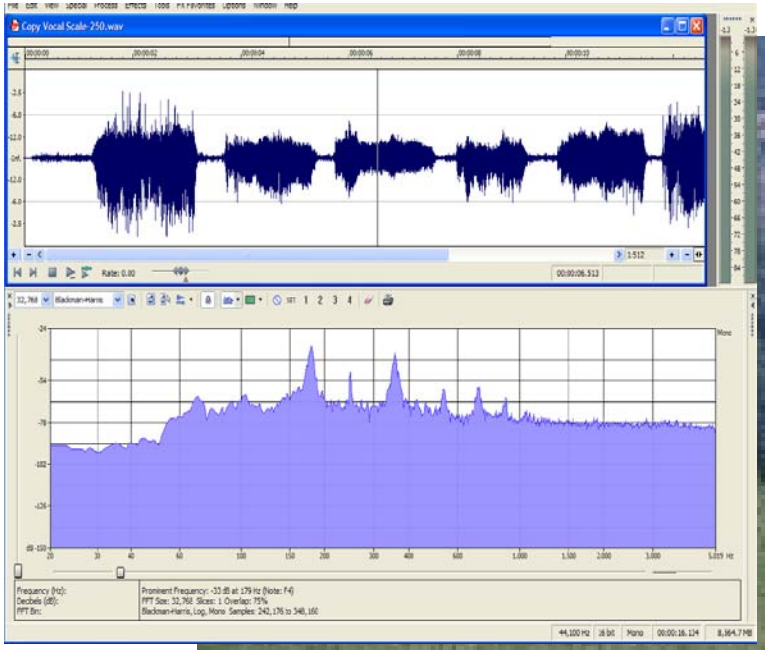


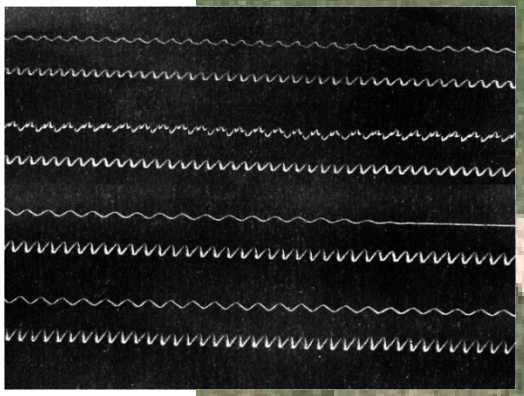




"Au Clair de la Lune" ["By the Light of the Moon"] sung;
"…the pitch is measured by the tuning fork of 500 simple vibrations per second which writes directly and simultaneously in interlinear space of the song"

Léon Scott 9 April 1860



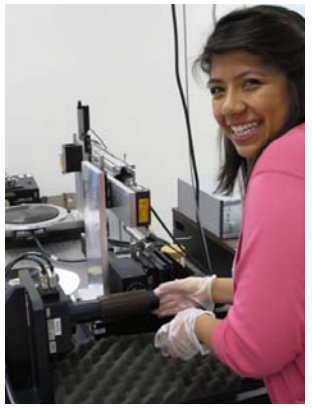


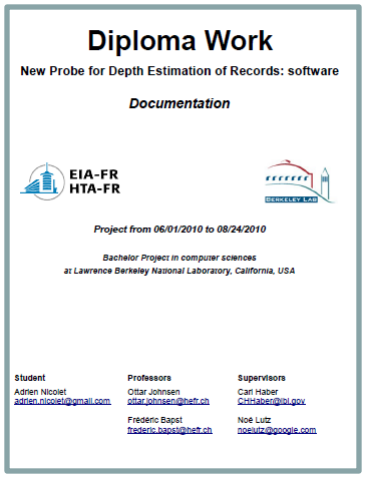
2nd example of a sound recording in history w/ 250 Hz tuning fork crosstalk

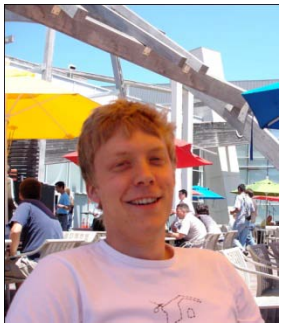
Do Re Mi Fa So La Ti# Do

Education and Training

- At Berkeley we have developed a significant student participation
- ~10 UCB undergraduate internships
- 8 thesis students (3 months FT) from the University of Applied Science EIF Fribourg, Switzerland
- Subjects: EE, CS, Physics
- Students have participated in measurements, data analysis, code development, algorithms
- Opportunity to expose students in engineering and physical science to problems in preservation/conservation







Past Student Projects

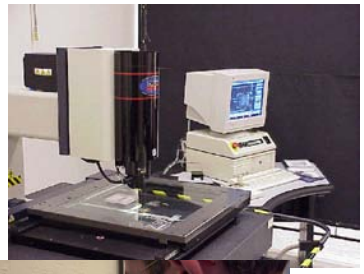
- Pilot scanning of wax cylinders, study of analysis parameters
- Application to plastic dictation belts
- Development of fast position finding on discs
- Study of interpolation algorithms
- Tools to measure signal and noise
- Effects of probe orientation on data quality
- Software to analyze groove shape in 3D
- Mechanical design of lightweight support frame

Future Student Participation

- Past students have been undergrad apprentices or visitors from Europe, this will continue
- Our funder(s) are supportive of more training, outreach, and student participation
  - We may be able to generate research support depending upon interest level, appropriateness etc.
  - Definite focus on science for preservation and increasing awareness of this aspect in academia and the national labs

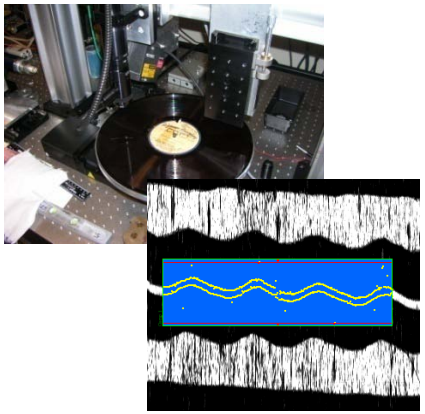
Optical Scanning: Technology Roadmap

Basic **2**D concept
demonstration
40 min / 1 sec

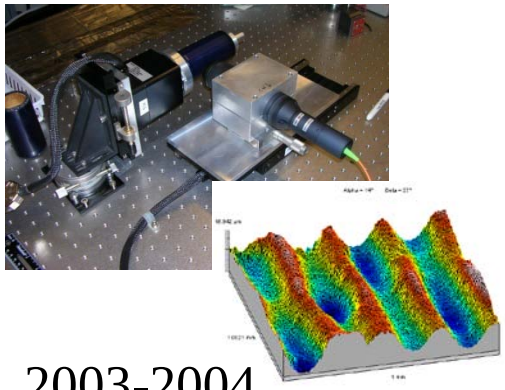


2002

2D disc R&D
4 sec / 1 sec



3D cylinder R&D
20 hr / 1 min

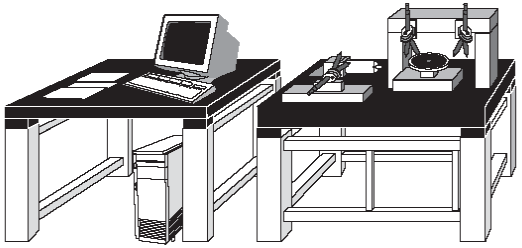


2003-2004

IRENE
System eval

2006-2007

3D System
10 min / 1 min



2008-20….

Production
mode discs

2009-10

Special
studies

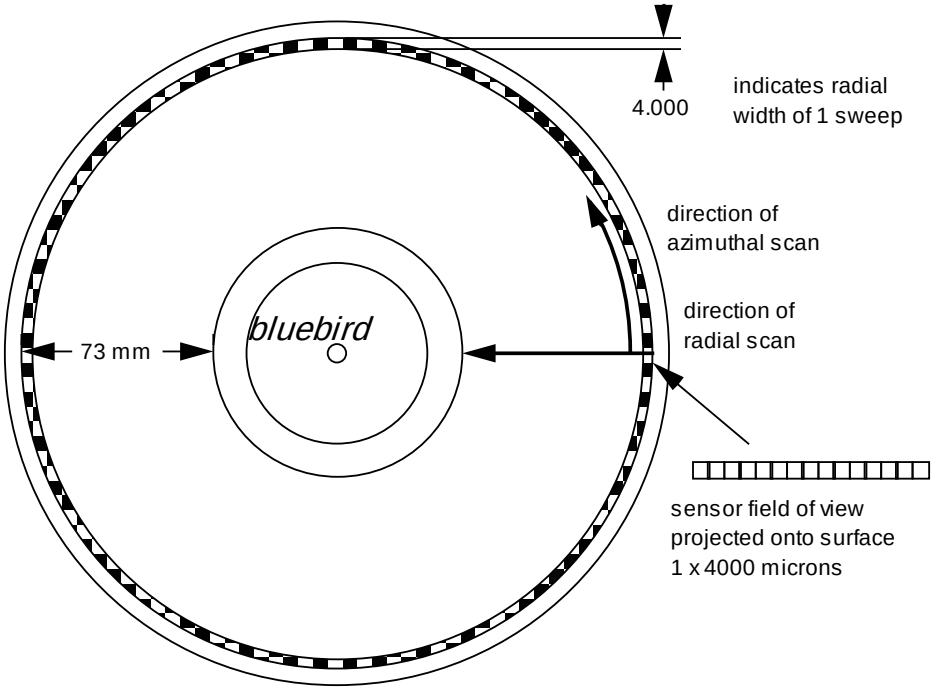
Web site: http://irene.lbl.gov/

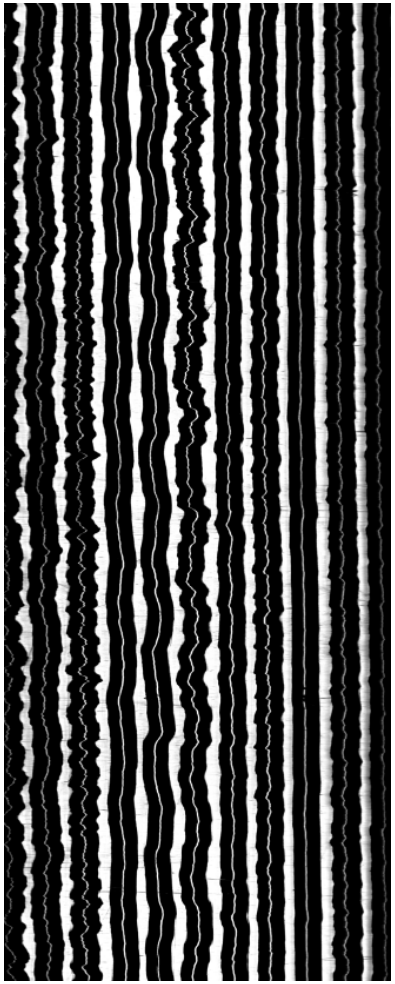
Backup Slides

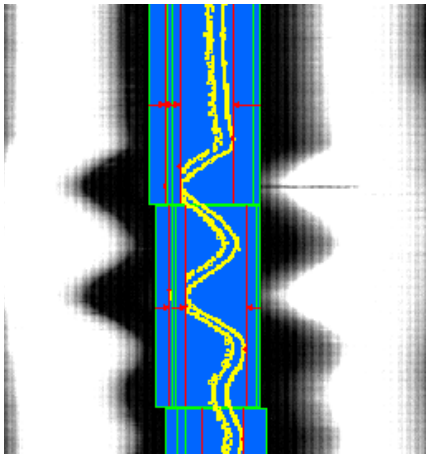
Milestones

- 2002: first demonstration of 2D method
- 2003: First 3D scans with single point probe
- 2004: Begin the development of a fast 2D scanner IRENE
  3D scans of dictabelts demonstrated
- 2006: IRENE installed at the Library of Congress for testing
- 2007: 3D scans of unique Jack London dictation recordings
  First 3D studies of Native American field recordings
- 2008: First high speed 3D probe acquired
  Restoration of Scott phonautograph, earliest sound recording using IRENE codes
  Begin to convert IRENE control system to express GUI and tools
- 2009: V2 3D probe acquired, enables efficient disc scanning for the first time
  Scan of Bishop Museum Kalakaua cylinder – nothing audible found
  IRENE moved to Packard Campus
- 2010: 3D installed at Library of Congress
  Systematic 3D study of F.Boas field recordings
  3D demonstrated on Galvanos
  3D disc measurement advances

Line Scanning: disc is in motion







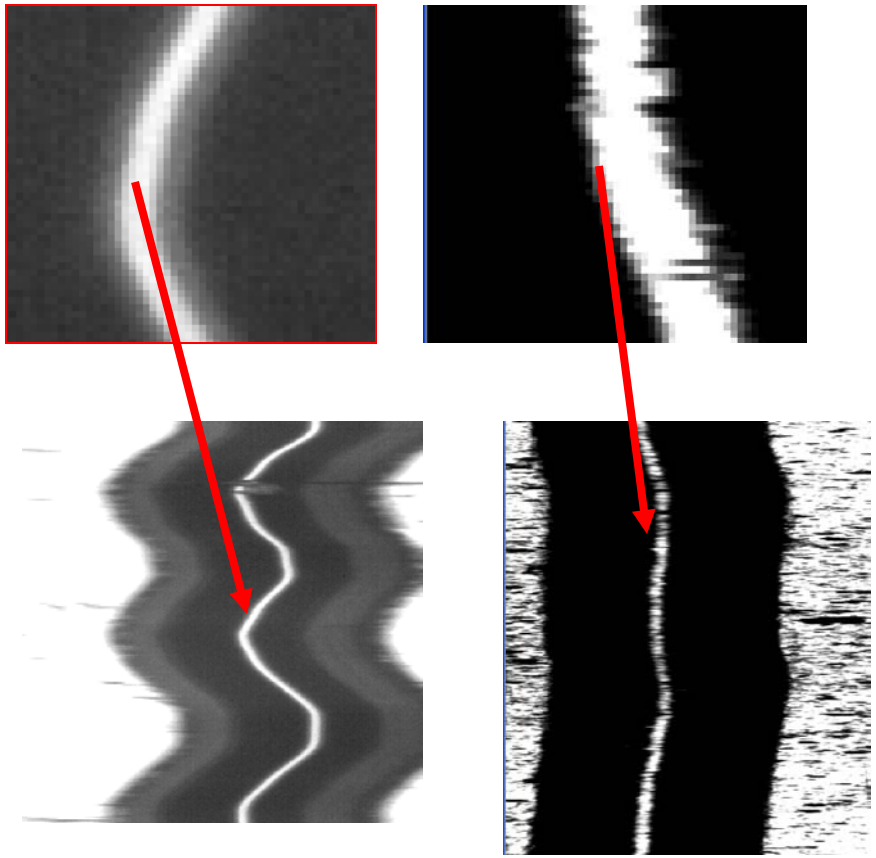
- 4000 pixels@~5x10^3 lines/s
- Requires bright illumination
- 7.6 x 10^5 lines/outer ring
  - **390 KHz max sampling**
- Scans @ a few x real time
- Scan time decreases linearly with sampling!!!.

Media Characteristics

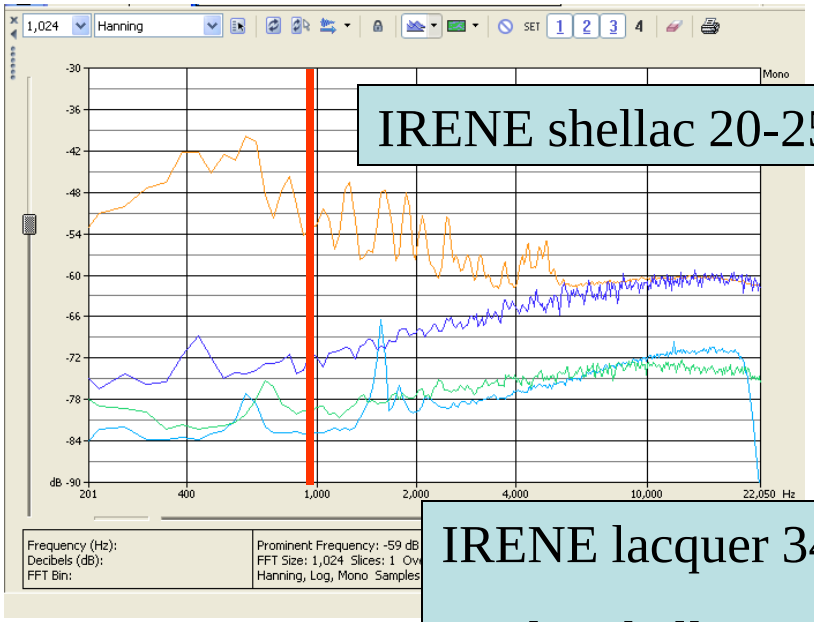
lacquer

shellac

Signal and noise @ 1 KHz



IRENE audio



IRENE shellac 20-25 dB

IRENE lacquer 34 dB

Stylus shellac 30 dB